

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jaye, Michael C.

Doan, Kim-Anh T.

Krawiec, John A.

Lynch, Kevin J.

Amin, Dilip V.

South, Victoria J.

Marchadier, Dawn

Maugeais, Cyrille

Rader, Daniel J.

(ii) TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR EFFECTING THE
LEVELS OF HIGH DENSITY LIPOPROTEIN (HDL) CHOLESTEROL AND
APOLIPOPROTEIN AI, VERY LOW DENSITY LIPOPROTEIN (VLDL)
CHOLESTEROL AND LOW DENSITY LIPOPROTEIN (LDL) CHOLESTEROL

(iii) NUMBER OF SEQUENCES: 31

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Synnestvedt & Lechner LLP

(B) STREET: Suite 2600 Aramark Tower, 1101 Market Street

(C) CITY: Philadelphia

(D) STATE: PA

(E) COUNTRY: USA

(F) ZIP: 19107

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-Windows 95

(D) SOFTWARE: Corel WordPerfect 8.0 converted to ASCII

(vi) CURRENT APPLICATION DATA:

60900-7042200

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Kelly, Ph.D., Patrick J.

(B) REGISTRATION NUMBER: 34,638

(C) REFERENCE/DOCKET NUMBER: 22,944 USA

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (215)923-4466

(B) TELEFAX: (215)923-2189

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 367 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 22..180

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAATTCGGCT TGATCAATCG C TTC AAA AAG GGG ATC TGT CTG AGC TGC CGC	51
Phe Lys Lys Gly Ile Cys Leu Ser Cys Arg	
1 5 10	
AAG AAC CGT TGT AAT AGC ATT GGC TAC AAT GCC AAG AAA ATG AGG AAC	99
Lys Asn Arg Cys Asn Ser Ile Gly Tyr Asn Ala Lys Lys Met Arg Asn	
15 20 25	
AAG AGG AAC AGC AAA ATG TAC CTA AAA ACC CGG GCA GGC ATG CCT TTC	147
Lys Arg Asn Ser Lys Met Tyr Leu Lys Thr Arg Ala Gly Met Pro Phe	
30 35 40	

AGA GGT AAC CTT CAG TCC CTG GAG TGT CCC TGA GGAAGGCCCT TAATACCTCC 200

Arg Gly Asn Leu Gln Ser Leu Glu Cys Pro *

45

50

TTCTTAATAC CATGCTGCAG AGCAGGGCAC ATCCTAGCCC AGGAGAAGTG GCCAGCACAA 260

TCCAATCAAA TCGTTGCAAA TCAGATTACA CTGTGCATGT CCTAGGAAAG GGAATCTTTA 320

CAAATAAAC AGTGTGGACC CCTCAAAAAA AAAAAAAGC CGAATTC 367

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Phe Lys Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Ser

1 5 10 15

Ile Gly Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys Met

20 25 30

Tyr Leu Lys Thr Arg Ala Gly Met Pro Phe Arg Gly Asn Leu Gln Ser

35 40 45

Leu Glu Cys Pro *

50

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1382 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

009220-1042250

(ix) **FEATURE:**

(B) LOCATION: 312..1370

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAATTCGGCT	TCTACTACTA	CTAGGCCACG	CGTCGCCTAG	TACGGGGGGG	GGGGGGGGGG	60										
TCAGCGAGTC	CTTGCCCTCCC	GGCGGCTCAG	GACGAGGGCA	GATCTCGTTC	TGGGGCAAGC	120										
CGTTGACACT	CGCTCCCTGC	CACCGCCCGG	GCTCCGTGCC	GCCAAGTTTT	CATTTTCCAC	180										
CTTCTCTGCC	TCCAGTCCCC	CAGCCCCTGG	CCGAGAGAAG	GGTCTTACCG	GCCGGGATTG	240										
CTGGAAACAC	CAAGAGGTGG	TTTTTGTTTT	TTAAACTTC	TGTTTCTTGG	GAGGGGGTGT	300										
GGCGGGGCAG	G	ATG	AGC	AAC	TCC	GTT	CCT	CTG	CTC	TGT	TTC	TGG	AGC	CTC	350	
Met Ser Asn Ser Val Pro Leu Leu Cys Phe Trp Ser Leu																
55					60					65						
TGC	TAT	TGC	TTT	GCT	GCG	GGG	AGC	CCC	GTA	CCT	TTT	GGT	CCA	GAG	GGA	398
Cys	Tyr	Cys	Phe	Ala	Ala	Gly	Ser	Pro	Val	Pro	Phe	Gly	Pro	Glu	Gly	
70					75					80						
CGG	CTG	GAA	GAT	AAG	CTC	CAC	AAA	CCC	AAA	GCT	ACA	CAG	ACT	GAG	GTC	446
Arg	Leu	Glu	Asp	Lys	Leu	His	Lys	Pro	Lys	Ala	Thr	Gln	Thr	Glu	Val	
85					90					95						
AAA	CCA	TCT	GTG	AGG	TTT	AAC	CTC	CGC	ACC	TCC	AAG	GAC	CCA	GAG	CAT	494
Lys	Pro	Ser	Val	Arg	Phe	Asn	Leu	Arg	Thr	Ser	Lys	Asp	Pro	Glu	His	
100					105					110						
GAA	GGA	TGC	TAC	CTC	TCC	GTC	GGC	CAC	AGC	CAG	CCC	TTA	GAA	GAC	TGC	542
Glu	Gly	Cys	Tyr	Leu	Ser	Val	Gly	His	Ser	Gln	Pro	Leu	Glu	Asp	Cys	
115					120					125					130	
AGT	TTC	AAC	ATG	ACA	GCT	AAA	ACC	TTT	TTC	ATC	ATT	CAC	GGA	TGG	ACG	590
Ser	Phe	Asn	Met	Thr	Ala	Lys	Thr	Phe	Phe	Ile	Ile	His	Gly	Trp	Thr	
135					140					145						

ATG	AGC	GGT	ATC	TTT	GAA	AAC	TGG	CTG	CAC	AAA	CTC	GTG	TCA	GCC	CTG	638
Met	Ser	Gly	Ile	Phe	Glu	Asn	Trp	Leu	His	Lys	Leu	Val	Ser	Ala	Leu	
			150					155					160			
CAC	ACA	AGA	GAG	AAA	GAC	GCC	AAT	GTA	GTT	GTG	GTT	GAC	TGG	CTC	CCC	686
His	Thr	Arg	Glu	Lys	Asp	Ala	Asn	Val	Val	Val	Val	Asp	Trp	Leu	Pro	
			165					170					175			
CTG	GCC	CAC	CAG	CTT	TAC	ACG	GAT	GCG	GTC	AAT	AAT	ACC	AGG	GTG	GTG	734
Leu	Ala	His	Gln	Leu	Tyr	Thr	Asp	Ala	Val	Asn	Asn	Thr	Arg	Val	Val	
			180					185					190			
GGA	CAC	AGC	ATT	GCC	AGG	ATG	CTC	GAC	TGG	CTG	CAG	GAG	AAG	GAC	GAT	782
Gly	His	Ser	Ile	Ala	Arg	Met	Leu	Asp	Trp	Leu	Gln	Glu	Lys	Asp	Asp	
195					200					205					210	
TTT	TCT	CTC	GGG	AAT	GTC	CAC	TTG	ATC	GGC	TAC	AGC	CTC	GGA	GCG	CAC	830
Phe	Ser	Leu	Gly	Asn	Val	His	Leu	Ile	Gly	Tyr	Ser	Leu	Gly	Ala	His	
				215					220					225		
GTG	GCC	GGG	TAT	GCA	GGC	AAC	TTC	GTG	AAA	GGA	ACG	GTG	GGC	CGA	ATC	878
Val	Ala	Gly	Tyr	Ala	Gly	Asn	Phe	Val	Lys	Gly	Thr	Val	Gly	Arg	Ile	
			230					235					240			
ACA	GGT	TTG	GAT	CCT	GCC	GGG	CCC	ATG	TTT	GAA	GGG	GCC	GAC	ATC	CAC	926
Thr	Gly	Leu	Asp	Pro	Ala	Gly	Pro	Met	Phe	Glu	Gly	Ala	Asp	Ile	His	
			245					250					255			
AAG	AGG	CTC	TCT	CCG	GAC	GAT	GCA	GAT	TTT	GTG	GAT	GTC	CTC	CAC	ACC	974
Lys	Arg	Leu	Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His	Thr	
			260					265					270			
TAC	ACG	CGT	TCC	TTC	GGC	TTG	AGC	ATT	GGT	ATT	CAG	ATG	CCT	GTG	GGC	1022
Tyr	Thr	Arg	Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly	
275					280					285					290	
CAC	ATT	GAC	ATC	TAC	CCC	AAT	GGG	GGT	GAC	TTC	CAG	CCA	GGC	TGT	GGA	1070
His	Ile	Asp	Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly	
				295						300					305	

CTC AAC GAT GTC TTG GGA TCA ATT GCA TAT GGA ACA ATC ACA GAG GTG	1118
Leu Asn Asp Val Leu Gly Ser Ile Ala Tyr Gly Thr Ile Thr Glu Val	
310 315 320	
GTA AAA TGT GAG CAT GAG CGA GCC GTC CAC CTC TTT GTT GAC TCT CTG	1166
Val Lys Cys Glu His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu	
325 330 335	
GTG AAT CAG GAC AAG CCG AGT TTT GCC TTC CAG TGC ACT GAC TCC AAT	1214
Val Asn Gln Asp Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn	
340 345 350	
CGC TTC AAA AAG GGG ATC TGT CTG AGC TGC CGC AAG AAC CGT TGT AAT	1262
Arg Phe Lys Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn	
355 360 365 370	
AGC ATT GGC TAC AAT GCC AAG AAA ATG AGG AAC AAG AGG AAC AGC AAA	1310
Ser Ile Gly Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys	
375 380 385	
ATG TAC CTA AAA ACC CGG GCA GGC ATG CCT TTC AGA GGT AAC CTT CAG	1358
Met Tyr Leu Lys Thr Arg Ala Gly Met Pro Phe Arg Gly Asn Leu Gln	
390 395 400	
TCC CTG GAG TGT CAAGCCGAAT TC	1382
Ser Leu Glu Cys	
405	

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 353 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

009220-1042250

Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly	His	Ile	Asp
225				230				235				240			
Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly	Leu	Asn	Asp
245				250				255							
Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val	Val	Lys	Cys
260				265				270							
Glu	His	Glu	Arg	Ala	Val	His	Leu	Phe	Val	Asp	Ser	Leu	Val	Asn	Gln
275				280				285							
Asp	Lys	Pro	Ser	Phe	Ala	Phe	Gln	Cys	Thr	Asp	Ser	Asn	Arg	Phe	Lys
290				295				300							
Lys	Gly	Ile	Cys	Leu	Ser	Cys	Arg	Lys	Asn	Arg	Cys	Asn	Ser	Ile	Gly
305				310				315				320			
Tyr	Asn	Ala	Lys	Lys	Met	Arg	Asn	Lys	Arg	Asn	Ser	Lys	Met	Tyr	Leu
325				330				335							
Lys	Thr	Arg	Ala	Gly	Met	Pro	Phe	Arg	Gly	Asn	Leu	Gln	Ser	Leu	Glu
340				345				350							
Cys															

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: nucleic acid

(D) TOPOLOGY: linear

(ix) **FEATURE:**

(B) LOCATION: 1..1065

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG	AGC	AAC	TCC	GTT	CCT	CTG	CTC	TGT	TTC	TGG	AGC	CTC	TGC	TAT	TGC	48
Met	Ser	Asn	Ser	Val	Pro	Leu	Leu	Cys	Phe	Trp	Ser	Leu	Cys	Tyr	Cys	
355				360				365								
TTT	GCT	GCG	GGG	AGC	CCC	GTA	CCT	TTT	GGT	CCA	GAG	GGA	CGG	CTG	GAA	96
Phe	Ala	Ala	Gly	Ser	Pro	Val	Pro	Phe	Gly	Pro	Glu	Gly	Arg	Leu	Glu	
370				375				380				385				
GAT	AAG	CTC	CAC	AAA	CCC	AAA	GCT	ACA	CAG	ACT	GAG	GTC	AAA	CCA	TCT	144
Asp	Lys	Leu	His	Lys	Pro	Lys	Ala	Thr	Gln	Thr	Glu	Val	Lys	Pro	Ser	
390				395				400								
GTG	AGG	TTT	AAC	CTC	CGC	ACC	TCC	AAG	GAC	CCA	GAG	CAT	GAA	GGA	TGC	192
Val	Arg	Phe	Asn	Leu	Arg	Thr	Ser	Lys	Asp	Pro	Glu	His	Glu	Gly	Cys	
405				410				415								
TAC	CTC	TCC	GTC	GGC	CAC	AGC	CAG	CCC	TTA	GAA	GAC	TGC	AGT	TTC	AAC	240
Tyr	Leu	Ser	Val	Gly	His	Ser	Gln	Pro	Leu	Glu	Asp	Cys	Ser	Phe	Asn	
420				425				430								
ATG	ACA	GCT	AAA	ACC	TTT	TTC	ATC	ATT	CAC	GGA	TGG	ACG	ATG	AGC	GGT	288
Met	Thr	Ala	Lys	Thr	Phe	Phe	Ile	Ile	His	Gly	Trp	Thr	Met	Ser	Gly	
435				440				445								
ATC	TTT	GAA	AAC	TGG	CTG	CAC	AAA	CTC	GTG	TCA	GCC	CTG	CAC	ACA	AGA	336
Ile	Phe	Glu	Asn	Trp	Leu	His	Lys	Leu	Val	Ser	Ala	Leu	His	Thr	Arg	
450				455				460				465				
GAG	AAA	GAC	GCC	AAT	GTA	GTT	GTG	GTT	GAC	TGG	CTC	CCC	CTG	GCC	CAC	384
Glu	Lys	Asp	Ala	Asn	Val	Val	Val	Val	Asp	Trp	Leu	Pro	Leu	Ala	His	
470				475				480								
CAG	CTT	TAC	ACG	GAT	GCG	GTC	AAT	AAT	ACC	AGG	GTG	GTG	GGA	CAC	AGC	432
Gln	Leu	Tyr	Thr	Asp	Ala	Val	Asn	Asn	Thr	Arg	Val	Val	Gly	His	Ser	
485				490				495								

ATT GCC AGG ATG CTC GAC TGG CTG CAG GAG AAG GAC GAT TTT TCT CTC	480
Ile Ala Arg Met Leu Asp Trp Leu Gln Glu Lys Asp Asp Phe Ser Leu	
500 505 510	
GGG AAT GTC CAC TTG ATC GGC TAC AGC CTC GGA GCG CAC GTG GCC GGG	528
Gly Asn Val His Leu Ile Gly Tyr Ser Leu Gly Ala His Val Ala Gly	
515 520 525	
TAT GCA GGC AAC TTC GTG AAA GGA ACG GTG GGC CGA ATC ACA GGT TTG	576
Tyr Ala Gly Asn Phe Val Lys Gly Thr Val Gly Arg Ile Thr Gly Leu	
530 535 540 545	
GAT CCT GCC GGG CCC ATG TTT GAA GGG GCC GAC ATC CAC AAG AGG CTC	624
Asp Pro Ala Gly Pro Met Phe Glu Gly Ala Asp Ile His Lys Arg Leu	
550 555 560	
TCT CCG GAC GAT GCA GAT TTT GTG GAT GTC CTC CAC ACC TAC ACG CGT	672
Ser Pro Asp Asp Ala Asp Phe Val Asp Val Leu His Thr Tyr Thr Arg	
565 570 575	
TCC TTC GGC TTG AGC ATT GGT ATT CAG ATG CCT GTG GGC CAC ATT GAC	720
Ser Phe Gly Leu Ser Ile Gly Ile Gln Met Pro Val Gly His Ile Asp	
580 585 590	
ATC TAC CCC AAT GGG GGT GAC TTC CAG CCA GGC TGT GGA CTC AAC GAT	768
Ile Tyr Pro Asn Gly Gly Asp Phe Gln Pro Gly Cys Gly Leu Asn Asp	
595 600 605	
GTC TTG GGA TCA ATT GCA TAT GGA ACA ATC ACA GAG GTG GTA AAA TGT	816
Val Leu Gly Ser Ile Ala Tyr Gly Thr Ile Thr Glu Val Val Lys Cys	
610 615 620 625	
GAG CAT GAG CGA GCC GTC CAC CTC TTT GTT GAC TCT CTG GTG AAT CAG	864
Glu His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu Val Asn Gln	
630 635 640	
GAC AAG CCG AGT TTT GCC TTC CAG TGC ACT GAC TCC AAT CGC TTC AAA	912
Asp Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn Arg Phe Lys	
645 650 655	

AAG GGG ATC TGT CTG AGC TGC CGC AAG AAC CGT TGT AAT AGC ATT GGC	960
Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Ser Ile Gly	
660 665 670	
TAC AAT GCC AAG AAA ATG AGG AAC AAG AGG AAC AGC AAA ATG TAC CTA	1008
Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys Met Tyr Leu	
675 680 685	
AAA ACC CGG GCA GGC ATG CCT TTC AGA GGT AAC CTT CAG TCC CTG GAG	1056
Lys Thr Arg Ala Gly Met Pro Phe Arg Gly Asn Leu Gln Ser Leu Glu	
690 695 700 705	
TGT CCC TGA	1065
Cys Pro *	

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: amino acid

(ii) MOLECULE TYPE: protein

Met	Ser	Asn	Ser	Val	Pro	Leu	Leu	Cys	Phe	Trp	Ser	Leu	Cys	Tyr	Cys
1				5					10					15	
Phe	Ala	Ala	Gly	Ser	Pro	Val	Pro	Phe	Gly	Pro	Glu	Gly	Arg	Leu	Glu
			20					25					30		
Asp	Lys	Leu	His	Lys	Pro	Lys	Ala	Thr	Gln	Thr	Glu	Val	Lys	Pro	Ser
		35					40					45			
Val	Arg	Phe	Asn	Leu	Arg	Thr	Ser	Lys	Asp	Pro	Glu	His	Glu	Gly	Cys
	50					55				60					
Tyr	Leu	Ser	Val	Gly	His	Ser	Gln	Pro	Leu	Glu	Asp	Cys	Ser	Phe	Asn
65				70					75					80	

Met	Thr	Ala	Lys	Thr	Phe	Phe	Ile	Ile	His	Gly	Trp	Thr	Met	Ser	Gly
				85					90					95	
Ile	Phe	Glu	Asn	Trp	Leu	His	Lys	Leu	Val	Ser	Ala	Leu	His	Thr	Arg
			100					105					110		
Glu	Lys	Asp	Ala	Asn	Val	Val	Val	Val	Asp	Trp	Leu	Pro	Leu	Ala	His
			115					120					125		
Gln	Leu	Tyr	Thr	Asp	Ala	Val	Asn	Asn	Thr	Arg	Val	Val	Gly	His	Ser
			130				135						140		
Ile	Ala	Arg	Met	Leu	Asp	Trp	Leu	Gln	Glu	Lys	Asp	Asp	Phe	Ser	Leu
145					150					155					160
Gly	Asn	Val	His	Leu	Ile	Gly	Tyr	Ser	Leu	Gly	Ala	His	Val	Ala	Gly
				165					170					175	
Tyr	Ala	Gly	Asn	Phe	Val	Lys	Gly	Thr	Val	Gly	Arg	Ile	Thr	Gly	Leu
			180						185					190	
Asp	Pro	Ala	Gly	Pro	Met	Phe	Glu	Gly	Ala	Asp	Ile	His	Lys	Arg	Leu
			195					200					205		
Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His	Thr	Tyr	Thr	Arg
			210				215						220		
Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly	His	Ile	Asp
225					230					235					240
Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly	Leu	Asn	Asp
				245					250					255	
Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val	Val	Lys	Cys
			260						265					270	
Glu	His	Glu	Arg	Ala	Val	His	Leu	Phe	Val	Asp	Ser	Leu	Val	Asn	Gln
			275					280						285	
Asp	Lys	Pro	Ser	Phe	Ala	Phe	Gln	Cys	Thr	Asp	Ser	Asn	Arg	Phe	Lys
			290				295					300			
Lys	Gly	Ile	Cys	Leu	Ser	Cys	Arg	Lys	Asn	Arg	Cys	Asn	Ser	Ile	Gly
305					310					315					320

355

380

CGG CTG GAA GAT AAG CTC CAC AAA CCC AAA GCT ACA CAG ACT GAG GTC	386
Arg Leu Glu Asp Lys Leu His Lys Pro Lys Ala Thr Gln Thr Glu Val	
385 390 395 400	
AAA CCA TCT GTG AGG TTT AAC CTC CGC ACC TCC AAG GAC CCA GAG CAT	434
Lys Pro Ser Val Arg Phe Asn Leu Arg Thr Ser Lys Asp Pro Glu His	
405 410 415	
GAA GGA TGC TAC CTC TCC GTC GGC CAC AGC CAG CCC TTA GAA GAC TGC	482
Glu Gly Cys Tyr Leu Ser Val Gly His Ser Gln Pro Leu Glu Asp Cys	
420 425 430	
AGT TTC AAC ATG ACA GCT AAA ACC TTT TTC ATC ATT CAC GGA TGG ACG	530
Ser Phe Asn Met Thr Ala Lys Thr Phe Phe Ile Ile His Gly Trp Thr	
435 440 445	
ATG AGC GGT ATC TTT GAA AAC TGG CTG CAC AAA CTC GTG TCA GCC CTG	578
Met Ser Gly Ile Phe Glu Asn Trp Leu His Lys Leu Val Ser Ala Leu	
450 455 460	
CAC ACA AGA GAG AAA GAC GCC AAT GTA GTT GTG GTT GAC TGG CTC CCC	626
His Thr Arg Glu Lys Asp Ala Asn Val Val Val Val Asp Trp Leu Pro	
465 470 475 480	
CTG GCC CAC CAG CTT TAC ACG GAT GCG GTC AAT AAT ACC AGG GTG GTG	674
Leu Ala His Gln Leu Tyr Thr Asp Ala Val Asn Asn Thr Arg Val Val	
485 490 495	
GGA CAC AGC ATT GCC AGG ATG CTC GAC TGG CTG CAG GAG AAG GAC GAT	722
Gly His Ser Ile Ala Arg Met Leu Asp Trp Leu Gln Glu Lys Asp Asp	
500 505 510	
TTT TCT CTC GGG AAT GTC CAC TTG ATC GGC TAC AGC CTC GGA GCG CAC	770
Phe Ser Leu Gly Asn Val His Leu Ile Gly Tyr Ser Leu Gly Ala His	
515 520 525	
GTG GCC GGG TAT GCA GGC AAC TTC GTG AAA GGA ACG GTG GGC CGA ATC	818
Val Ala Gly Tyr Ala Gly Asn Phe Val Lys Gly Thr Val Gly Arg Ile	
530 535 540	

665220 T042260

ACA	GGT	TTG	GAT	CCT	GCC	GGG	CCC	ATG	TTT	GAA	GGG	GCC	GAC	ATC	CAC		866
Thr	Gly	Leu	Asp	Pro	Ala	Gly	Pro	Met	Phe	Glu	Gly	Ala	Asp	Ile	His		
545					550					555					560		
AAG	AGG	CTC	TCT	CCG	GAC	GAT	GCA	GAT	TTT	GTG	GAT	GTC	CTC	CAC	ACC		914
Lys	Arg	Leu	Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His	Thr		
				565					570					575			
TAC	ACG	CGT	TCC	TTC	GGC	TTG	AGC	ATT	GGT	ATT	CAG	ATG	CCT	GTG	GGC		962
Tyr	Thr	Arg	Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly		
			580						585					590			
CAC	ATT	GAC	ATC	TAC	CCC	AAT	GGG	GGT	GAC	TTC	CAG	CCA	GGC	TGT	GGA		1010
His	Ile	Asp	Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly		
			595				600							605			
CTC	AAC	GAT	GTC	TTG	GGA	TCA	ATT	GCA	TAT	GGA	ACA	ATC	ACA	GAG	GTG		1058
Leu	Asn	Asp	Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val		
			610				615							620			
GTA	AAA	TGT	GAG	CAT	GAG	CGA	GCC	GTC	CAC	CTC	TTT	GTT	GAC	TCT	CTG		1106
Val	Lys	Cys	Glu	His	Glu	Arg	Ala	Val	His	Leu	Phe	Val	Asp	Ser	Leu		
625					630					635					640		
GTG	AAT	CAG	GAC	AAG	CCG	AGT	TTT	GCC	TTC	CAG	TGC	ACT	GAC	TCC	AAT		1154
Val	Asn	Gln	Asp	Lys	Pro	Ser	Phe	Ala	Phe	Gln	Cys	Thr	Asp	Ser	Asn		
				645					650					655			
CGC	TTC	AAA	AAG	GGG	ATC	TGT	CTG	AGC	TGC	CGC	AAG	AAC	CGT	TGT	AAT		1202
Arg	Phe	Lys	Lys	Gly	Ile	Cys	Leu	Ser	Cys	Arg	Lys	Asn	Arg	Cys	Asn		
			660						665					670			
AGC	ATT	GGC	TAC	AAT	GCC	AAG	AAA	ATG	AGG	AAC	AAG	AGG	AAC	AGC	AAA		1250
Ser	Ile	Gly	Tyr	Asn	Ala	Lys	Lys	Met	Arg	Asn	Lys	Arg	Asn	Ser	Lys		
			675						680					685			
ATG	TAC	CTA	AAA	ACC	CGG	GCA	GGC	ATG	CCT	TTC	AGA	GTT	TAC	CAT	TAT		1298
Met	Tyr	Leu	Lys	Thr	Arg	Ala	Gly	Met	Pro	Phe	Arg	Val	Tyr	His	Tyr		
690					695									700			

CAG ATG AAA ATC CAT GTC TTC AGT TAC AAG AAC ATG GGA GAA ATT GAG	1346
Gln Met Lys Ile His Val Phe Ser Tyr Lys Asn Met Gly Glu Ile Glu	
705 710 715 720	
CCC ACC TTT TAC GTC ACC CTT TAT GGC ACT AAT GCA GAT TCC CAG ACT	1394
Pro Thr Phe Tyr Val Thr Leu Tyr Gly Thr Asn Ala Asp Ser Gln Thr	
725 730 735	
CTG CCA CTG GAA ATA GTG GAG CGG ATC GAG CAG AAT GCC ACC AAC ACC	1442
Leu Pro Leu Glu Ile Val Glu Arg Ile Glu Gln Asn Ala Thr Asn Thr	
740 745 750	
TTC CTG GTC TAC ACC GAG GAG GAC TTG GGA GAC CTC TTG AAG ATC CAG	1490
Phe Leu Val Tyr Thr Glu Glu Asp Leu Gly Asp Leu Leu Lys Ile Gln	
755 760 765	
CTC ACC TGG GAG GGG GCC TCT CAG TCT TGG TAC AAC CTG TGG AAG GAG	1538
Leu Thr Trp Glu Gly Ala Ser Gln Ser Trp Tyr Asn Leu Trp Lys Glu	
770 775 780	
TTT CGC AGC TAC CTG TCT CAA CCC CGC AAC CCC GGA CGG GAG CTG AAT	1586
Phe Arg Ser Tyr Leu Ser Gln Pro Arg Asn Pro Gly Arg Glu Leu Asn	
785 790 795 800	
ATC AGG CGC ATC CGG GTG AAG TCT GGG GAA ACC CAG CGG AAA CTG ACA	1634
Ile Arg Arg Ile Arg Val Lys Ser Gly Glu Thr Gln Arg Lys Leu Thr	
805 810 815	
TTT TGT ACA GAA GAC CCT GAG AAC ACC AGC ATA TCC CCA GGC CGG GAG	1682
Phe Cys Thr Glu Asp Pro Glu Asn Thr Ser Ile Ser Pro Gly Arg Glu	
820 825 830	
CTC TGG TTT CGC AAG TGT CGG GAT GGC TGG AGG ATG AAA AAC GAA ACC	1730
Leu Trp Phe Arg Lys Cys Arg Asp Gly Trp Arg Met Lys Asn Glu Thr	
835 840 845	
AGT CCC ACT GTG GAG CTT CCC TGA GGGTGCCCGG GCAAGTCTTG CCAGCAAGGC	1784
Ser Pro Thr Val Glu Leu Pro *	
850 855	

00320: T042260

AGCAAGACTT CCTGCTATCC AAGCCCATGG AGGAAAGTTA CTGCTGAGGA CCCACCCAAT	1844
GGAAGGATTC TTCTCAGCCT TGACCCTGGA GCACTGGGAA CAACTGGTCT CCTGTGATGG	1904
CTGGGACTCC TCGCGGGAGG GGACTGCGCT GCTATAGCTC TTGCTGCCTC TCTTGAATAG	1964
CTCTAACTCC AAACCTCTGT CCACACCTCC AGAGCACCAA GTCCAGATTT GTGTGTAAGC	2024
AGCTGGGTGC CTGGGGCCTC TCGTGACAC TGGATTGGTT TCTCAGTTGC TGGGCGAGCC	2084
TGTACTCTGC CTGACGAGGA ACGCTGGCTC CGAAGAGGCC CTGTGTAGAA GGCTGTCAGC	2144
TGCTCAGCCT GCTTTGAGCC TCAGTGAGAA GTCCTTCCGA CAGGAGCTGA CTCATGTCAG	2204
GATGGCAGGC CTGGTATCTT GCTCGGGCCC TGGCTGTTGG GGTTCATG GGTGCACTG	2264
ACCATACTGC TTACGTCTTA GCCATTCCGT CCTGCTCCCC AGCTCACTCT CTGAAGCACA	2324
CATCATTGGC TTTCTATTT TTCTGTTCAT TTTTAAATTG AGCAAATGTC TATTGAACAC	2384
TTAAAATTAA TTAGAATGTG GTAATGGACA TATTACTGAG CCTCTCCATT TGGAACCCAG	2444
TGGAGTTGGG ATTTCTAGAC CCTCTTTCTG TTTGGATGGT GTATGTGTAT ATGCATGGGG	2504
AAAGGCACCT GGGGCCTGGG GGAGGCTATA GGATATAAGC AGTCGACGCG GCCGCGAATT	2564
C	2565

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 501 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Ser	Asn	Ser	Val	Pro	Leu	Leu	Cys	Phe	Trp	Ser	Leu	Cys	Tyr	Cys
1				5					10					15	
Phe	Ala	Ala	Gly	Ser	Pro	Val	Pro	Phe	Gly	Pro	Glu	Gly	Arg	Leu	Glu
				20					25					30	
Asp	Lys	Leu	His	Lys	Pro	Lys	Ala	Thr	Gln	Thr	Glu	Val	Lys	Pro	Ser
				35					40					45	
Val	Arg	Phe	Asn	Leu	Arg	Thr	Ser	Lys	Asp	Pro	Glu	His	Glu	Gly	Cys
				50					55					60	

Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Ser Ile Gly
 305 310 315 320
 Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys Met Tyr Leu
 325 330 335
 Lys Thr Arg Ala Gly Met Pro Phe Arg Val Tyr His Tyr Gln Met Lys
 340 345 350
 Ile His Val Phe Ser Tyr Lys Asn Met Gly Glu Ile Glu Pro Thr Phe
 355 360 365
 Tyr Val Thr Leu Tyr Gly Thr Asn Ala Asp Ser Gln Thr Leu Pro Leu
 370 375 380
 Glu Ile Val Glu Arg Ile Glu Gln Asn Ala Thr Asn Thr Phe Leu Val
 385 390 395 400
 Tyr Thr Glu Glu Asp Leu Gly Asp Leu Leu Lys Ile Gln Leu Thr Trp
 405 410 415
 Glu Gly Ala Ser Gln Ser Trp Tyr Asn Leu Trp Lys Glu Phe Arg Ser
 420 425 430
 Tyr Leu Ser Gln Pro Arg Asn Pro Gly Arg Glu Leu Asn Ile Arg Arg
 435 440 445
 Ile Arg Val Lys Ser Gly Glu Thr Gln Arg Lys Leu Thr Phe Cys Thr
 450 455 460
 Glu Asp Pro Glu Asn Thr Ser Ile Ser Pro Gly Arg Glu Leu Trp Phe
 465 470 475 480
 Arg Lys Cys Arg Asp Gly Trp Arg Met Lys Asn Glu Thr Ser Pro Thr
 485 490 495
 Val Glu Leu Pro *
 500

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1035 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1035

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG AGC AAC TCC GTT CCT CTG CTC TGT TTC TGG AGC CTC TGC TAT TGC	48
Met Ser Asn Ser Val Pro Leu Leu Cys Phe Trp Ser Leu Cys Tyr Cys	
505 510 515	
TTT GCT GCG GGG AGC CCC GTA CCT TTT GGT CCA GAG GGA CGG CTG GAA	96
Phe Ala Ala Gly Ser Pro Val Pro Phe Gly Pro Glu Gly Arg Leu Glu	
520 525 530	
GAT AAG CTC CAC AAA CCC AAA GCT ACA CAG ACT GAG GTC AAA CCA TCT	144
Asp Lys Leu His Lys Pro Lys Ala Thr Gln Thr Glu Val Lys Pro Ser	
535 540 545	
GTG AGG TTT AAC CTC CGC ACC TCC AAG GAC CCA GAG CAT GAA GGA TGC	192
Val Arg Phe Asn Leu Arg Thr Ser Lys Asp Pro Glu His Glu Gly Cys	
550 555 560 565	
TAC CTC TCC GTC GGC CAC AGC CAG CCC TTA GAA GAC TGC AGT TTC AAC	240
Tyr Leu Ser Val Gly His Ser Gln Pro Leu Glu Asp Cys Ser Phe Asn	
570 575 580	
ATG ACA GCT AAA ACC TTT TTC ATC ATT CAC GGA TGG ACG ATG AGC GGT	288
Met Thr Ala Lys Thr Phe Phe Ile Ile His Gly Trp Thr Met Ser Gly	
585 590 595	
ATC TTT GAA AAC TGG CTG CAC AAA CTC GTG TCA GCC CTG CAC ACA AGA	336
Ile Phe Glu Asn Trp Leu His Lys Leu Val Ser Ala Leu His Thr Arg	
600 605 610	

GAG	AAA	GAC	GCC	AAT	GTA	GTT	GTG	GTT	GAC	TGG	CTC	CCC	CTG	GCC	CAC	384
Glu	Lys	Asp	Ala	Asn	Val	Val	Val	Val	Asp	Trp	Leu	Pro	Leu	Ala	His	
615							620					625				
CAG	CTT	TAC	ACG	GAT	GCG	GTC	AAT	AAT	ACC	AGG	GTG	GTG	GGA	CAC	AGC	432
Gln	Leu	Tyr	Thr	Asp	Ala	Val	Asn	Asn	Thr	Arg	Val	Val	Gly	His	Ser	
630						635					640				645	
ATT	GCC	AGG	ATG	CTC	GAC	TGG	CTG	CAG	GAG	AAG	GAC	GAT	TTT	TCT	CTC	480
Ile	Ala	Arg	Met	Leu	Asp	Trp	Leu	Gln	Glu	Lys	Asp	Asp	Phe	Ser	Leu	
				650					655					660		
GAG	AAT	GTC	CAC	TTG	ATC	GGC	TAC	AGC	CTC	GGA	GCG	CAC	GTG	GCC	GGG	528
Gly	Asn	Val	His	Leu	Ile	Gly	Tyr	Ser	Leu	Gly	Ala	His	Val	Ala	Gly	
			665					670					675			
TAT	GCA	GGC	AAC	TTC	GTG	AAA	GGA	ACG	GTG	GGC	CGA	ATC	ACA	GGT	TTG	576
Tyr	Ala	Gly	Asn	Phe	Val	Lys	Gly	Thr	Val	Gly	Arg	Ile	Thr	Gly	Leu	
	680						685					690				
GAT	CCT	GCC	GGG	CCC	ATG	TTT	GAA	GGG	GCC	GAC	ATC	CAC	AAG	AGG	CTC	624
Asp	Pro	Ala	Gly	Pro	Met	Phe	Glu	Gly	Ala	Asp	Ile	His	Lys	Arg	Leu	
	695						700					705				
TCT	CCG	GAC	GAT	GCA	GAT	TTT	GTG	GAT	GTC	CTC	CAC	ACC	TAC	ACG	CGT	672
Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His	Thr	Tyr	Thr	Arg	
710						715				720				725		
TCC	TTC	GGC	TTG	AGC	ATT	GGT	ATT	CAG	ATG	CCT	GTG	GGC	CAC	ATT	GAC	720
Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly	His	Ile	Asp	
				730						735				740		
ATC	TAC	CCC	AAT	GGG	GGT	GAC	TTC	CAG	CCA	GGC	TGT	GGA	CTC	AAC	GAT	768
Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly	Leu	Asn	Asp	
			745					750					755			
GTC	TTG	GGA	TCA	ATT	GCA	TAT	GGA	ACA	ATC	ACA	GAG	GTG	GTA	AAA	TGT	816
Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val	Val	Lys	Cys	
	760						765						770			

(2) INFORMATION FOR SEQ ID NO:10:

(A) LENGTH: 345 amino acids

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Ser	Asn	Ser	Val	Pro	Leu	Leu	Cys	Phe	Trp	Ser	Leu	Cys	Tyr	Cys
1				5					10					15	
Phe	Ala	Ala	Gly	Ser	Pro	Val	Pro	Phe	Gly	Pro	Glu	Gly	Arg	Leu	Glu
			20					25					30		
Asp	Lys	Leu	His	Lys	Pro	Lys	Ala	Thr	Gln	Thr	Glu	Val	Lys	Pro	Ser
			35					40					45		

Val	Arg	Phe	Asn	Leu	Arg	Thr	Ser	Lys	Asp	Pro	Glu	His	Glu	Gly	Cys
50				55				60							
Tyr	Leu	Ser	Val	Gly	His	Ser	Gln	Pro	Leu	Glu	Asp	Cys	Ser	Phe	Asn
65				70				75				80			
Met	Thr	Ala	Lys	Thr	Phe	Phe	Ile	Ile	His	Gly	Trp	Thr	Met	Ser	Gly
85				90				95							
Ile	Phe	Glu	Asn	Trp	Leu	His	Lys	Leu	Val	Ser	Ala	Leu	His	Thr	Arg
100				105				110							
Glu	Lys	Asp	Ala	Asn	Val	Val	Val	Val	Asp	Trp	Leu	Pro	Leu	Ala	His
115				120				125							
Gln	Leu	Tyr	Thr	Asp	Ala	Val	Asn	Asn	Thr	Arg	Val	Val	Gly	His	Ser
130				135				140							
Ile	Ala	Arg	Met	Leu	Asp	Trp	Leu	Gln	Glu	Lys	Asp	Asp	Phe	Ser	Leu
145				150				155				160			
Gly	Asn	Val	His	Leu	Ile	Gly	Tyr	Ser	Leu	Gly	Ala	His	Val	Ala	Gly
165				170				175							
Tyr	Ala	Gly	Asn	Phe	Val	Lys	Gly	Thr	Val	Gly	Arg	Ile	Thr	Gly	Leu
180				185				190							
Asp	Pro	Ala	Gly	Pro	Met	Phe	Glu	Gly	Ala	Asp	Ile	His	Lys	Arg	Leu
195				200				205							
Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His	Thr	Tyr	Thr	Arg
210				215				220							
Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly	His	Ile	Asp
225				230				235				240			
Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly	Leu	Asn	Asp
245				250				255							
Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val	Val	Lys	Cys
260				265				270							
Glu	His	Glu	Arg	Ala	Val	His	Leu	Phe	Val	Asp	Ser	Leu	Val	Asn	Gln
275				280				285							

Asp Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn Arg Phe Lys
 290 295 300
 Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Ser Ile Gly
 305 310 315 320
 Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys Met Tyr Leu
 325 330 335
 Lys Thr Arg Ala Gly Met Pro Phe Arg
 340 345

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 225 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..225

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTG GGA TCC ATC GCC TAT GGC ACG ATC GCG GAG GTG GTG AAG TGC GAG	48
Leu Gly Ser Ile Ala Tyr Gly Thr Ile Ala Glu Val Val Lys Cys Glu	
350 355 360	
CAT GAG CGG GCC GTG CAT CTC TTT GTG GAC TCC CTG GTG AAC CAG GAC	96
His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu Val Asn Gln Asp	
365 370 375	
AAG CCG AGC TTT GCC TTC CAG TGC ACA GAC TCC AAC CGC TTC AAA AAA	144
Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn Arg Phe Lys Lys	
380 385 390	
GGG ATC TGT CTC AGC TGC CGG AAG AAC CGC TGT AAC GGC ATC GGC TAC	192

66552207042250

Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Gly Ile Gly Tyr

395

400

405

AAT GCT AAG AAG ACG AGG AAT AAG AGG AAC ACC

225

Asn Ala Lys Lys Thr Arg Asn Lys Arg Asn Thr

410

415

420

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Leu Gly Ser Ile Ala Tyr Gly Thr Ile Ala Glu Val Val Lys Cys Glu

1

5

10

15

His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu Val Asn Gln Asp

20

25

30

Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn Arg Phe Lys Lys

35

40

45

Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Gly Ile Gly Tyr

50

55

60

Asn Ala Lys Lys Thr Arg Asn Lys Arg Asn Thr

65

70

75

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 472 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

6656207042250

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met	Glu	Ser	Lys	Ala	Leu	Leu	Val	Leu	Thr	Leu	Ala	Val	Trp	Leu	Gln
1			5					10						15	
Ser	Leu	Thr	Ala	Ser	Arg	Gly	Gly	Val	Ala	Ala	Ala	Asp	Gln	Arg	Arg
			20					25					30		
Asp	Phe	Ile	Asp	Ile	Glu	Ser	Lys	Phe	Ala	Leu	Arg	Thr	Pro	Glu	Asp
			35					40					45		
Thr	Ala	Glu	Asp	Thr	Cys	His	Leu	Ile	Pro	Gly	Val	Ala	Glu	Ser	Val
			50				55					60			
Ala	Thr	Cys	His	Phe	Asn	His	Ser	Ser	Lys	Thr	Phe	Met	Val	Ile	His
65					70					75					80
Gly	Trp	Thr	Val	Thr	Gly	Met	Tyr	Glu	Ser	Trp	Val	Pro	Lys	Leu	Val
				85						90				95	
Ala	Ala	Leu	Tyr	Lys	Arg	Glu	Pro	Asp	Ser	Asn	Val	Ile	Val	Val	Asp
				100					105				110		
Trp	Leu	Ser	Arg	Ala	Gln	Glu	His	Tyr	Pro	Val	Ser	Ala	Gly	Tyr	Thr
				115					120				125		
Lys	Val	Gly	Gln	Asp	Val	Ala	Arg	Phe	Ile	Asn	Trp	Met	Glu	Glu	Glu
						130					135			140	
Phe	Asn	Tyr	Pro	Leu	Asp	Asn	Val	His	Leu	Leu	Gly	Tyr	Ser	Leu	Gly
145					150						155				160
Ala	His	Ala	Ala	Gly	Ile	Ala	Gly	Ser	Leu	Thr	Asn	Lys	Lys	Val	Asn
					165					170				175	
Arg	Ile	Thr	Gly	Leu	Asp	Pro	Ala	Gly	Pro	Asn	Phe	Glu	Tyr	Ala	Glu
						180				185				190	
Ala	Pro	Arg	Leu	Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His
						195				200			205		
Thr	Phe	Thr	Arg	Gly	Ser	Pro	Gly	Arg	Ser	Ile	Gly	Ile	Gln	Lys	Pro
						210						215			220

465

(2) INFORMATION FOR SEQ ID NO:14:

(A) LENGTH: 499 amino acids

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

1 5 10 15

20 25 30

35 40 45

50 55 60

65 70 75 80

85 90 95

100 105 110

115 120 125

130 135 140

Ala Ala Leu Leu Arg Trp Leu Glu Glu Ser Val Gln Leu Ser Arg Ser
145 150 155 160
His Val His Leu Ile Gly Tyr Ser Leu Gly Ala His Val Ser Gly Phe
165 170 175
Ala Gly Ser Ser Ile Gly Gly Thr His Lys Ile Gly Arg Ile Thr Gly
180 185 190
Leu Asp Ala Ala Gly Pro Leu Phe Glu Gly Ser Ala Pro Ser Asn Arg
195 200 205
Leu Ser Pro Asp Asp Ala Asn Phe Val Asp Ala Ile His Thr Phe Thr
210 215 220
Arg Glu His Met Gly Leu Ser Val Gly Ile Lys Gln Pro Ile Gly His
225 230 235 240
Tyr Asp Phe Tyr Pro Asn Gly Gly Ser Phe Gln Pro Gly Cys His Phe
245 250 255
Leu Glu Leu Tyr Arg His Ile Ala Gln His Gly Phe Asn Ala Ile Thr
260 265 270
Gln Thr Ile Lys Cys Ser His Glu Arg Ser Val His Leu Phe Ile Asp
275 280 285
Ser Leu Leu His Ala Gly Thr Gln Ser Met Ala Tyr Pro Cys Gly Asp
290 295 300
Met Asn Ser Phe Ser Gln Gly Leu Cys Leu Ser Cys Lys Lys Gly Arg
305 310 315 320
Cys Asn Thr Leu Gly Tyr His Val Arg Gln Glu Pro Arg Ser Lys Ser
325 330 335
Lys Arg Leu Phe Leu Val Thr Arg Ala Gln Ser Pro Phe Lys Val Tyr
340 345 350
His Tyr Gln Leu Lys Ile Gln Phe Ile Asn Gln Thr Glu Thr Pro Ile
355 360 365
Gln Thr Thr Phe Thr Met Ser Leu Leu Gly Thr Lys Glu Lys Met Gln
370 375 380

00920 104200

-

Lys	Asp	Val	Asn	Thr	Arg	Phe	Leu	Leu	Tyr	Thr	Asn	Glu	Asn	Pro	Asn		
50						55						60					
Asn	Phe	Gln	Glu	Val	Ala	Ala	Asp	Ser	Ser	Ser	Ile	Ser	Gly	Ser	Asn		
65						70						75			80		
Phe	Lys	Thr	Asn	Arg	Lys	Thr	Arg	Phe	Ile	Ile	His	Gly	Phe	Ile	Asp		
			85						90						95		
Lys	Gly	Glu	Glu	Asn	Trp	Leu	Ala	Asn	Val	Cys	Lys	Asn	Leu	Phe	Lys		
			100						105						110		
Val	Glu	Ser	Val	Asn	Cys	Ile	Cys	Val	Asp	Trp	Lys	Gly	Gly	Ser	Arg		
			115						120						125		
Thr	Gly	Tyr	Thr	Gln	Ala	Ser	Gln	Asn	Ile	Arg	Ile	Val	Gly	Ala	Glu		
130						135						140					
Val	Ala	Tyr	Phe	Val	Glu	Phe	Leu	Gln	Ser	Ala	Phe	Gly	Tyr	Ser	Pro		
145						150						155			160		
Ser	Asn	Val	His	Val	Ile	Gly	His	Ser	Leu	Gly	Ala	His	Ala	Ala	Gly		
			165						170						175		
Glu	Ala	Gly	Arg	Arg	Thr	Asn	Gly	Thr	Ile	Gly	Arg	Ile	Thr	Gly	Leu		
			180						185						190		
Asp	Pro	Ala	Glu	Pro	Cys	Phe	Gln	Gly	Thr	Pro	Glu	Leu	Val	Arg	Leu		
195						200						205					
Asp	Pro	Ser	Asp	Ala	Lys	Phe	Val	Asp	Val	Ile	His	Thr	Asp	Gly	Ala		
210						215						220					
Pro	Ile	Val	Pro	Asn	Leu	Gly	Phe	Gly	Met	Ser	Gln	Val	Val	Gly	His		
225						230						235			240		
Leu	Asp	Phe	Phe	Pro	Asn	Gly	Gly	Val	Glu	Met	Pro	Gly	Cys	Lys	Lys		
			245						250						255		
Asn	Ile	Leu	Ser	Gln	Ile	Val	Asp	Ile	Asp	Gly	Ile	Trp	Glu	Gly	Thr		
			260						265						270		
Arg	Asp	Phe	Ala	Ala	Cys	Asn	His	Leu	Arg	Ser	Tyr	Lys	Tyr	Tyr	Thr		
275						280						285					

Asp Ser Ile Val Asn Pro Asp Gly Phe Ala Gly Phe Pro Cys Ala Ser
290 295 300
Tyr Asn Val Phe Thr Ala Asn Lys Cys Phe Pro Cys Pro Ser Gly Gly
305 310 315 320
Cys Pro Gln Met Gly His Tyr Ala Asp Arg Tyr Pro Gly Lys Thr Asn
325 330 335
Asp Val Gly Gln Lys Phe Tyr Leu Asp Thr Gly Asp Ala Ser Asn Phe
340 345 350
Ala Arg Trp Arg Tyr Lys Val Ser Val Thr Leu Ser Gly Lys Lys Val
355 360 365
Thr Gly His Ile Leu Val Ser Leu Phe Gly Asn Lys Gly Asn Ser Lys
370 375 380
Gln Tyr Glu Ile Phe Lys Gly Thr Leu Lys Pro Asp Ser Thr His Ser
385 390 395 400
Asn Glu Phe Asp Ser Asp Val Asp Val Gly Asp Leu Gln Met Val Lys
405 410 415
Phe Ile Trp Tyr Asn Asn Val Ile Asn Pro Thr Leu Pro Arg Val Gly
420 425 430
Ala Ser Lys Ile Ile Val Glu Thr Asn Val Gly Lys Gln Phe Asn Phe
435 440 445
Cys Ser Pro Glu Thr Val Arg Glu Glu Val Leu Leu Thr Leu Thr Pro
450 455 460
Cys
465

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Cys

(i) SEQUENCE CHARACTERISTICS:

(D) TOPOLOGY: linear

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

13

(i) SEQUENCE CHARACTERISTICS:

(D) TOPOLOGY: linear

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

10

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TAGGACATGC ACAGTGAAT CTG

23

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GATTGTGCTG GCCACTTCTC

20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCACTCCAG GGACTGAAG

19

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 36

(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 37

(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 41

(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 42

(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 46

66520-1042250

(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 47

(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CUACUACUAC UAGGCCACGC GTCGACTAGT ACGGGNNGGG NNGGGNNG

48

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CACACACAGG CCACGCGTCG ACTAGTAC

28

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ACCACCATGG AGAGCAAAGC CCTG

24

663220 1042250

CCGGGTGGGT AGGTACATTT TG

22

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GGGGGTGACT TCCAGCCAGG CTGTG

25

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AACTCTGAAA GGCATGCCTG CCCGG

25

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TGAAGGTCGG AGTCAACGGA TTGGT

26

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CATGTGGGCC ATGAGGTCCA CCAC

24

663220 1042250